

EXCERPT OF REFERENCE 2

Japanese Patent Laid-open No. H11-161682

Laid-open on June 18, 1999

Japanese Patent Application No. H10-139539

Filed on May 21, 1999

Title of the Invention: Information Search Device and Method

Description of the Invention:

The present invention relates to an information search device and method for conducting information search on the basis of designated words and idioms without increasing any user load.

Fig. 1 shows a block diagram of an example of an information search device according to the invention, comprising an input unit 101, a search control unit 102, a parsing unit 103, a database search unit 104, an analysis knowledge 105, a database 106 and an output unit 107. The information search device is configured to submit search results which are expected to be requested by a user.

Output unit 107 can provide search results in the form of such multimedia information as photograph, drawing, sound, image and text. When any search character stream is designated, a type of information which can be displayed by the search character stream can be submitted to a user by means of a pop-up menu to give the user a chance to select information to be displayed. If a user wishes to obtain detailed information about the word "director", the user drags the character stream "director", as shown by a meshed portion in Fig. 30. In response, search control unit 102 causes displayable multimedia information to be submitted to the user in a pop-up menu. Then, the user selects a desired type of information from the pop-up menu. Alternatively, instead of selecting a type of information to conduct a data search, it is possible to conduct a data search before submitting displayable information to display search results corresponding to the type which the user has designated.

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-161682

(43) 公開日 平成11年(1999) 6月18日

(51) Int.Cl.*

識別記号

F I

G 0 6 F 17/30

G 0 6 F 15/403

3 3 0 C

17/27

15/38

E

17/28

S

審査請求 未請求 請求項の数13 O L (全 17 頁)

(21) 出願番号 特願平10-139539

(22) 出願日 平成10年(1998) 5月21日

(31) 優先権主張番号 特願平9-263606

(32) 優先日 平9 (1997) 9月29日

(33) 優先権主張国 日本 (J P)

(71) 出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72) 発明者 酒匂 孝之

神奈川県川崎市幸区小向東芝町1番地 株

式会社東芝研究開発センター内

(72) 発明者 木下 聡

神奈川県川崎市幸区小向東芝町1番地 株

式会社東芝研究開発センター内

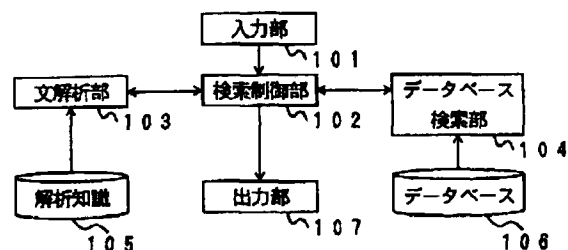
(74) 代理人 弁理士 鈴江 武彦 (外6名)

(54) 【発明の名称】 情報検索装置、情報検索方法及び記録媒体

(57) 【要約】

【課題】 ユーザの求めている適切な検索結果を自動的に呈示することにより、これまで必要であったユーザが検索結果を絞り込む操作を軽減することを可能とした情報検索装置を提供すること。

【解決手段】 文書中の単語や熟語の文字データを指定し、指定された文字データに基づいて所定のデータベースから対象となるデータを検索し、該文字データを含む少なくとも1文を切り出し該文の解析を行ない、該解析結果に基づいてデータベースから検索された検索結果の中から特定のデータを選択する。また、文書中の単語や熟語の文字データを指定し、指定された文字データを含む少なくとも1文を切り出し、該文の解析を行なってキーワードを抽出し、抽出されたキーワードから検索条件式を作成し、作成された検索条件に基づいて所定のデータベースから対象となるデータを検索する。



【特許請求の範囲】

【請求項1】文書中の単語あるいは熟語の文字データを指定する手段と、

少なくとも前記文書中における前記指定された文字データと所定の関係にある部分に基づいて、1または複数のキーワードを求め手段と、

少なくとも前記指定された文字データおよび前記求められたキーワードに基づいて、所定のデータベースに対する検索結果として提示すべきデータを決定する手段とを備えたことを特徴とする情報検索装置。

【請求項2】文書中の単語あるいは熟語の文字データを指定する手段と、

この手段により指定された文字データに基づいて、所定のデータベースから対象となるデータを検索する手段と、

前記文字データを含む少なくとも1文を切り出し、該文の解析を行なう手段と、

この手段による文の解析結果に基づいて、前記データベースから検索された検索結果の中から特定のデータを選択する手段とを備えたことを特徴とする情報検索装置。

【請求項3】前記文の解析結果は、該文の形態素解析結果によって抽出されたキーワードであることを特徴とする請求項2に記載の情報検索装置。

【請求項4】前記文の解析結果は、該文の形態素解析結果によって抽出されたキーワードのうち、距離的に近いもののみを選択したものであることを特徴とする請求項2に記載の情報検索装置。

【請求項5】文書中の単語あるいは熟語の文字データを指定する手段と、

この手段により指定された文字データを含む少なくとも1文を切り出し、該文の解析を行なってキーワードを抽出する手段と、

この手段により抽出されたキーワードから検索条件式を作成する手段と、

この手段により作成された検索条件に基づいて、所定のデータベースから対象となるデータを検索する手段とを備えたことを特徴とする情報検索装置。

【請求項6】抽出されたキーワードから検索条件を作成する際に、抽出されたキーワードのうち、距離的に近いもののみを用いて検索条件を作成することを特徴とする請求項5に記載の情報検索装置。

【請求項7】抽出されたキーワードから所定のデータベースを用いて他の関連するキーワードを抽出し、それらのキーワードのうち少なくとも1つを用いて特定のデータを選択するを特徴とする請求項2ないし6のいずれか1項に記載の情報検索装置。

【請求項8】文書中の単語あるいは熟語の文字データを指定し、

少なくとも前記文書中における前記指定された文字データと所定の関係にある部分に基づいて、1または複数の

キーワードを求め、

少なくとも前記指定された文字データおよび前記求められたキーワードに基づいて、所定のデータベースに対する検索結果として提示すべきデータを決定することを特徴とする情報検索方法。

【請求項9】文書中の単語あるいは熟語の文字データを指定し、

指定された文字データに基づいて、所定のデータベースから対象となるデータを検索し、

10 前記文字データを含む少なくとも1文を切り出し、該文の解析を行い、

解析された文の解析結果に基づいて、前記データベースから検索された検索結果の中から特定のデータを選択することを特徴とする情報検索方法。

【請求項10】文書中の単語あるいは熟語の文字データを指定し、

指定された文字データを含む少なくとも1文を切り出し、該文の解析を行なってキーワードを抽出し、

抽出されたキーワードから検索条件式を作成し、

20 作成された検索条件に基づいて、所定のデータベースから対象となるデータを検索することを特徴とする情報検索方法。

【請求項11】コンピュータに、文書中の単語あるいは熟語の文字データの入力を受けさせ、少なくとも前記文書中における前記指定された文字データと所定の関係にある部分に基づいて、1または複数のキーワードを求めさせ、少なくとも前記指定された文字データおよび前記求められたキーワードに基づいて、所定のデータベースに対する検索結果として提示すべきデータを決定させるためのプログラムを記録したコンピュータ読取り可能な記録媒体。

【請求項12】文書中で指定された単語あるいは熟語の文字データの人力を受け、

該文字データに基づいて、所定のデータベースから対象となるデータを検索し、

該文字データを含む少なくとも1文を切り出し、該文の解析を行い、

該文の解析結果に基づいて、前記データベースから検索された検索結果の中から特定のデータを選択する機能を備えたことを特徴とする情報検索のための記録媒体。

【請求項13】文書中で指定された単語あるいは熟語の文字データの人力を受け、

該文字データを含む少なくとも1文を切り出し、該文の解析を行なってキーワードを抽出し、

抽出されたキーワードから検索条件式を作成し、

作成された検索条件に基づいて、所定のデータベースから対象となるデータを検索する機能を備えたことを特徴とする情報検索のための記録媒体。

【発明の詳細な説明】

50 【0001】

【発明の属する技術分野】本発明は、指定された単語あるいは熟語に基づいて情報検索を行う情報検索装置、情報検索方法及び記録媒体に関する。

【0002】

【従来の技術】計算機の小型化やインターネットの普及に伴い、文書の作成や閲覧という操作を計算機上で行うという状況が増えてきており、文書の推敲や辞書引きを計算機上で行うためのソフトウェアも既に実用化されている。

【0003】例えば、辞書検索ソフトでは辞書データをCD-ROM等に格納しておき、ユーザが文書を作成しているときや文書を読んでいるときに、詳細情報が欲しい文字列を検索文字列として入力することにより、CD-ROM等に蓄積されている情報を参照することが可能になっている。

【0004】

【発明が解決しようとする課題】しかしながら、従来のデータ検索ソフトではCD-ROM等に蓄積されているデータを検索抽出し、抽出に成功したデータを全てユーザに呈示するという手法しかとられていなかった。これでは、ユーザは欲しい情報を得るのに、さらにそれらのデータの中から、自分が必要な情報を手作業で探さなくてはならず、ユーザにとっての負荷は大きなものであった。

【0005】本発明は、上記事情を考慮してなされたもので、これまで必要であった検索結果からさらに必要な情報をユーザが絞り込むという負荷を軽減させる情報検索装置、情報検索方法及び記録媒体を提供することを目的とする。

【0006】

【課題を解決するための手段】本発明（請求項1）に係る情報検索装置は、文書中の単語あるいは熟語の文字データを指定する手段と、少なくとも前記文書中における前記指定された文字データと所定の関係にある部分に基づいて、1または複数のキーワードを求め手段と、少なくとも前記指定された文字データおよび前記求められたキーワードに基づいて、所定のデータベースに対する検索結果として提示すべきデータを決定する手段とを備えたことを特徴とする。

【0007】本発明（請求項2）に係る情報検索装置は、文書中の単語あるいは熟語の文字データを指定する手段と、この手段により指定された文字データに基づいて、所定のデータベースから対象となるデータを検索する手段と、前記文字データを含む少なくとも1文を切り出し、該文の解析を行なう手段と、この手段による文の解析結果に基づいて、前記データベースから検索された検索結果の中から特定のデータを選択する手段とを備えたことを特徴とする。

【0008】好ましくは、前記文の解析結果は、該文の形態素解析結果によって抽出されたキーワードであるよ

うにしてもよい。

【0009】好ましくは、前記文の解析結果は、該文の形態素解析結果によって抽出されたキーワードのうち、距離的に近いもののみを選択したものであるようにしてもよい。

【0010】本発明（請求項5）に係る情報検索装置は、文書中の単語あるいは熟語の文字データを指定する手段と、この手段により指定された文字データを含む少なくとも1文を切り出し、該文の解析を行なってキーワードを抽出する手段と、この手段により抽出されたキーワードから検索条件式を作成する手段と、この手段により作成された検索条件に基づいて、所定のデータベースから対象となるデータを検索する手段とを備えたことを特徴とする。

【0011】好ましくは、抽出されたキーワードから検索条件を作成する際に、抽出されたキーワードのうち、距離的に近いもののみを用いて検索条件を作成するようにしてもよい。

【0012】好ましくは、抽出されたキーワードから所定のデータベースを用いて他の関連するキーワードを抽出し、それらのキーワードのうち少なくとも1つを用いて特定のデータを選択するようにしてもよい。

【0013】本発明（請求項8）に係る情報検索方法は、文書中の単語あるいは熟語の文字データを指定し、少なくとも前記文書中における前記指定された文字データと所定の関係にある部分に基づいて、1または複数のキーワードを求め、少なくとも前記指定された文字データおよび前記求められたキーワードに基づいて、所定のデータベースに対する検索結果として提示すべきデータを決定することを特徴とする。

【0014】本発明（請求項9）に係る情報検索方法は、文書中の単語あるいは熟語の文字データを指定し、指定された文字データに基づいて、所定のデータベースから対象となるデータを検索し、前記文字データを含む少なくとも1文を切り出し、該文の解析を行い、解析された文の解析結果に基づいて、前記データベースから検索された検索結果の中から特定のデータを選択することを特徴とする。

【0015】本発明（請求項10）に係る情報検索方法は、文書中の単語あるいは熟語の文字データを指定し、指定された文字データを含む少なくとも1文を切り出し、該文の解析を行なってキーワードを抽出し、抽出されたキーワードから検索条件式を作成し、作成された検索条件に基づいて、所定のデータベースから対象となるデータを検索することを特徴とする。

【0016】本発明（請求項11）は、コンピュータに、文書中の単語あるいは熟語の文字データの入力を受けさせ、少なくとも前記文書中における前記指定された文字データと所定の関係にある部分に基づいて、1または複数のキーワードを求めさせ、少なくとも前記指定さ

れた文字データおよび前記求められたキーワードに基づいて、所定のデータベースに対する検索結果として提示すべきデータを決定させるためのプログラムを記録したコンピュータ読取り可能な記録媒体を要旨とする。

【0017】本発明（請求項12）は、文書中で指定された単語あるいは熟語の文字データの入力を受け、該文字データに基づいて、所定のデータベースから対象となるデータを検索し、該文字データを含む少なくとも1文を切り出し、該文の解析を行い、該文の解析結果に基づいて、前記データベースから検索された検索結果の中から特定のデータを選択する機能を備えたことを特徴とする情報検索のための記録媒体を要旨とする。

【0018】本発明（請求項13）は、文書中で指定された単語あるいは熟語の文字データの入力を受け、該文字データを含む少なくとも1文を切り出し、該文の解析を行なってキーワードを抽出し、抽出されたキーワードから検索条件式を作成し、作成された検索条件に基づいて、所定のデータベースから対象となるデータを検索する機能を備えたことを特徴とする情報検索のための記録媒体を要旨とする。

【0019】なお、装置に係る本発明は方法に係る発明としても成立し、方法に係る本発明は装置に係る発明としても成立する。

【0020】また、装置または方法に係る本発明は、コンピュータに当該発明に相当する手順を実行させるための（あるいはコンピュータを当該発明に相当する手段として機能させるための、あるいはコンピュータに当該発明に相当する機能を実現させるための）プログラムを記録したコンピュータ読取り可能な記録媒体としても成立する。

【0021】本発明では、ユーザが指定した文字列の前後の情報などを自動的に抽出し、それらの抽出した情報を使って、検索結果を表示する際に表示させる情報を絞り込むことができる。

【0022】例えば、文書中の関連する、あるいは詳細な情報が知りたい単語、熟語等の文字データを指定すると、指定された文字データに基づいて、所定のデータベースから対象となるデータを検索し、文字データを含む少なくとも1文を切り出し、該文の解析を行い、解析された文の解析結果に基づいて、データベースから検索された検索結果の中から特定のデータを選択する。あるいは、例えば、指定された文字データを含む少なくとも1文を切り出し、該文の解析を行なってキーワードを抽出し、抽出されたキーワードから検索条件式を作成し、作成された検索条件に基づいて、所定のデータベースから対象となるデータを検索する。

【0023】本発明によれば、ユーザは必要な情報を効率良く得ることができ、文章の理解を早めることに効果がある。つまり、ユーザは必要な情報を自分で絞り込むことなく情報が呈示されるため、ユーザの負荷が軽減さ

れるのである。

【0024】

【発明の実施の形態】以下、図面を参照しながら発明の実施の形態を説明する。

【0025】（第1の実施形態）本発明の第1の実施形態について説明する。

【0026】図1に本実施形態に係る情報検索装置の構成例を示す。図1に示されるように、本情報検索装置は、入力部101、検索制御部102、文解析部103、データベース検索部104、解析知識105、データベース106、出力部107を備えている。

【0027】入力部101は、文書を入力するためのものであるとともに、また入力された文書中の検索すべき単語あるいは熟語の指示や検索された表示情報の選択の指示等を入力するためのものである。

【0028】文書の入力には、例えば、キーボード（キー入力）、記録媒体駆動装置（磁気ディスクや磁気テープあるいは光ディスクなどの記録媒体からの読み込み）、OCR（光学読取り）、ネットワーク接続装置（通信による取得）などのいずれか、もしくは複数種類のものを使用する。なお、OCRを使用する場合には、読み取った文字画像を文字認識してコード列を生成する機能を用いる。

【0029】ユーザからの指示の入力には、例えば、キーボード、マウスなどを使用する。

【0030】入力部101から入力された文書データは、検索制御部102のメモリ（図示せず）に一旦記憶され、その後、入力部101を介して表示中の文書におけるある文字データ（単語や熟語）の検索が指示されると、文書データの所定の部分（例えば、指示された文字データを含む1文あるいは該1文を少なくとも含む複数文）が文解析部103へ渡される。

【0031】入力部101によって検索が指示された文字データは、検索制御部102を介してデータベース検索部104に入力される。

【0032】文解析部103は、検索制御部102の制御の下、該検索制御部102から文字データを受け取って解析知識105を利用して文解析を行い解析結果を出力する。文解析の方法としては、形態素解析、構文解析、意味解析など、一般的な手法を用いることができる。

【0033】データベース検索部104は、検索制御部102の制御の下、該検索制御部102から文字データを受け取ってデータベース106を参照して検索結果を出力する。

【0034】検索制御部102は、文解析部103による文解析結果とデータベース検索部104によるデータベース106の検索結果とを利用して、ユーザが求めているデータを選択し、出力部107に渡す。

【0035】出力部107は、入力文書や検索結果をユ

ーザに呈示するためのものである。

【0036】出力部107としては、光学ディスプレイやプリンタなどのいずれか、もしくは複数種類のものを使用する。

【0037】以下、具体例を用いながら本実施形態をより詳しく説明する。

【0038】ここでは文書は既に入力され表示されているものとする。ユーザは入力部101によって、文書内にある詳細情報が得たい文字列（以下、検索文字列と呼ぶ）を指示する。

【0039】具体例として、図2に例示するように、「…彼はこの映画の監督である。…。…」といった文書が表示された表示画面においてユーザがマウス等の入力部101を利用して「監督」という文字列部分をドラッグし、「監督」に関する詳細な情報を得るために指示したとする。

【0040】検索制御部102は、位置情報等を利用してユーザの指示した検索文字列を含む1文を切り出し文解析部103に送る。また、検索文字列をデータベース検索部104に送る。

【0041】図2の例では、検索文字列「監督」を含む1文、「彼はこの映画の監督である。」が切り出され、文解析部103に送られる。また、検索文字列「監督」はデータベース検索部104に送られる。

【0042】文解析部103は、検索制御部102から入力された1文に対して、解析知識105を利用して解析処理を行い、解析結果を検索制御部102に渡す。ここで言う解析処理とは、形態素解析、構文解析、意味解析等の全て、もしくはその一部を指す。ここでは、一例として、形態素解析を利用した場合について説明する。図3に、上記の1文すなわち「彼はこの映画の監督である。」に対して形態素解析を行った場合の解析結果の一例を示す。

【0043】データベース検索部104は、検索制御部102から入力された検索文字列「監督」を使ってデータベース106の中から対象となるデータを取り出し、検索制御部102に渡す。

【0044】例えば、データベース106が図4に例示する国語辞書のようなものであるとすると、検索制御部102から渡された「監督」という検索文字列を利用して、その語義である「人の上に立って指図すること、人、機関」「映画・演劇などで演出する人」等が取り出される。

【0045】ここで、図4のデータの“< >”で囲まれた部分はタグである。例えば、“<見出し語>”から、“<\見出し語>”で挟まれた部分が辞書の見出し語となる。この例では辞書の見出し語は“監督”となる。次に見出し語に対する品詞情報が記述されている。ここでは、名詞、動詞「スル」と結び付いてサ行変格活用を行なう動詞として用いられる名詞であることを示

す。次に見出し語の語義が記述されている。ここでは語義が3つあり、それぞれが、“語義部1”から“語義部3”に記述されている。

【0046】なお、本実施形態においては、文解析部103による処理とデータベース検索部104による処理は、これらを同時に実行してもよいし、いずれか一方を先に他方を後に実行してもよい。

【0047】文解析部103からの解析結果およびデータベース検索部104からの検索結果は、検索制御部102に返され、ユーザに表示する内容が選択される。

【0048】ここで、図5に本実施形態における検索制御部102の構成例を示す。図5に示されるように、この検索制御部102は、制御記憶部601、キーワード抽出部602および表示内容選択部603から構成される。

【0049】制御記憶部601は、入力部101から入力された文書を記憶したり、ユーザからの検索指示を受け取りデータベース検索部104に送ったり、1文を切り出し文解析部103に送るといった、検索全体に関わる制御を行う。

【0050】キーワード抽出部602は、文解析部103から渡された解析結果の中から、表示内容を選択するのに有効なキーワードを抽出し、キーワードリストを作成して制御記憶部601に渡す。制御記憶部601は、キーワードリストとデータベース検索部104からの検索結果（図4参照）を表示内容選択部603に渡す。

【0051】図6に、キーワード抽出部602のキーワードリスト作成処理の手順の一例を示す。

【0052】まず、変数nに1をセットする（ステップS702）。この値は検索文字列（図3の場合、「監督」）からの距離を表す。ここで言う距離とは、検索文字列から何語離れているかという値である。

【0053】形態素解析結果は、一語一語キーワードになるかどうか判断される（ステップS703、S704）。ここで言うキーワードは、検索結果を絞り込むためのもので、名詞であるか、固有名詞であるかといった品詞情報や、コンピュータ用語、経済用語などといった分野情報等をキーとして抽出された用語を指す。ここではキーワード抽出のためのキーを名詞とする。キーワードとなる語はキーワードリストに保管される（ステップS705）。

【0054】さらに、変数nがスコープを越えていないか判断する（ステップS706）。ここで言うスコープとは、検索文字列から何語離れた語までをキーワード抽出の対象とするかを表わす値とする。例えば、スコープを5とすると、検索文字列の前後5単語からキーワードを抽出することになる。本実施形態ではスコープを3とする。スコープを越えていない場合は、nを1増やし（ステップS707）、再度キーワードの検索を行う。越えている場合は、スコープ内の全ての語が検索された

ので処理を終了し(ステップS708)、処理結果であるキーワードリストは表示内容選択部603に渡される。キーワードが抽出されなかった場合は、空のキーワードリストが表示内容選択部603に渡される。

【0055】以上の処理により、例えば図3に示す形態素解析結果が入力された場合には、「映画」がキーワードとして抽出され、表示内容選択部603に渡される。

【0056】表示内容選択部603は、キーワード抽出部602から渡されたキーワードリストとデータベース検索部104から渡された検索結果とを受け取り、表示内容を選択するために、検索結果に対して評価点を計算し、その点数の最も高いものを出力部107に渡す。

【0057】図7に表示内容の選択手法としての手順の一例を示す。この例では各語義文におけるキーワードの出現回数を調べ、その回数の最も多いものをユーザに表示する。本具体例では、表示内容選択部603は、図4で表わされる検索結果とキーワードリスト「映画」とを制御記憶部601より受け取ったとする。

【0058】まず、キーワード数テーブルを初期化する(ステップS902)。キーワード数テーブルは、語義データの中に表われたキーワードの数を記憶させておくテーブルで、語義データの数だけ用意し、一対一に対応させる。図4で表わされる検索結果を受け取った場合、この検索結果には語義が3つ含まれるので、図8に示すように、3つのキーワード数テーブルが用意される。

【0059】次に、変数nに1をセットする(ステップS903)。この値は語義データのエン트리番号を表わす。

【0060】語義データは、キーワードリスト内の文字列があるかどうか検索される(ステップS904、S905)。キーワードがある場合は、キーワードの個数をカウントし、語義データと一対一に対応したキーワード数テーブルに代入する(ステップS906)。

【0061】次の語義データがあるかどうか判断される(ステップS907)、ある場合はnを1増やし(ステップS908)、再度、語義データ内のキーワード検索を行う。キーワードがない場合は、処理を終了し(ステップS909)、キーワード数テーブルを利用して、キーワードが最も多く表われた語義が出力部107に送られ、辞書検索結果が表示される。

【0062】本具体例では、語義データの一番目にはキーワードである「映画」が存在しないため、図9に示すように、一番目のキーワード数テーブルは0である。二番目の語義データの中には「映画」が2語存在するため、二番目のキーワード数テーブルに2が代入される(ステップS906)。以下同様にして、各語義データについてキーワード数を求める。この結果、最も適切な語義として図4における二番目の語義データが選ばれる。

【0063】そして、例えば、図10に示すように、出

力部107により辞書検索結果が表示される。

【0064】なお、図10の表示例では検索文字列に最も関連するデータのみを表示しているが、辞書検索結果の表示の形態にはこの他にも種々のバリエーションが考えられる。

【0065】例えば、ポップアップしている表示画面を、マウス等を使ってクリックする方法により、他のデータ(例えば図4における一番目や三番目の語義データ)を表示するようにしてもよい。次候補は、次にキーワードの出現回数が多いもの、エン트리番号の小さいものといったような手法を単独もしくは組み合わせることによって、各候補に優先順位をつけ、それにより選択することができる。

【0066】図11に、上記具体例においてユーザが次候補の表示を指示した場合の表示例を示す。なお、図11の例ではエン트리番号の小さいものを優先するという優先順位を使用している。

【0067】また、例えば、図12に示すように、一つの検索結果を選択して表示するのではなく、全ての検索結果を優先順位で並び替えて表示するようにしてもよい。

【0068】なお、本実施形態では、キーワードの検索範囲として、検索対象語から離れている語数によるスコープを設定し、そのスコープの範囲内の語としたが、同じ内容を表す要約文、文章の先頭、切り出した一文内、一文を切出しその文中に見つからない場合はその前後の文を検索する、といった方法を用いてもよい。

【0069】また、データベース106として種々のものを使用することが可能であり、例えばデータベース106を類義語辞書のようなものとする、図13に例示するように、「監督」に関する類義語を選び、表示させることもできる。

【0070】以上のように本実施形態によれば、検索結果をそのまま提示するのではなく、よりユーザが望むと期待されるものだけを選択して提示し、あるいは検索結果をすべて提示する場合であってもよりユーザが望むと期待されるものがどれであるかを知らしめる形で提示するなどの提示方法をとることができる。この結果、ユーザは必要な情報を効率良く得ることができる。つまり、ユーザは必要な情報を自分で絞り込むことなく情報が呈示されるため、文章の理解が早まり、従来に比較してユーザの負荷を軽減することができる。

【0071】(第2の実施形態)次に、本発明の第2の実施形態について説明する。

【0072】第1の実施形態では、文解析部103の解析結果として形態素解析を利用する場合について説明したが、本実施形態では、文解析部103の解析結果として構文解析を利用する場合について説明する。

【0073】なお、以下では、第1の実施形態と相違する点を中心に説明する。

【0074】まず、ユーザによる検索指示の入力に応じた処理、データベース検索部104による処理は、第1の実施形態と同様である。

【0075】ここで、図14に、文解析部103の解析結果として形態素解析・構文解析を利用した場合の出力結果の一例を示す。文解析部103の解析結果が、図14で表わすような構文解析結果（概念依存構造データ）の場合、検索制御部102のキーワード抽出部602を変更することにより構文解析結果を利用した場合の情報検索装置を実現できる。

【0076】図15に構文解析結果を利用する場合における検索制御部102内のキーワード抽出部602の処理手順の一例を示す。

【0077】まず、変数nに1をセットする（ステップS802）。この値は検索文字列を示すノードからの距離を表す。ここで言う距離とは、検索文字列から何回リンクを辿ったかという値である。すなわち、図14の「監督」のノードから見ると「映画」「である」のノードは距離が1で「彼」のノードは距離が2となる。

【0078】構文解析結果は、一語一語キーワードになるかどうか判断される（ステップS803、S804）。ここで言うキーワードは、検索結果を絞り込むためのもので、名詞であるか、固有名詞であるかといった品詞情報や、コンピュータ用語、経済用語などといった分野情報等をキーとして抽出された用語を指す。ここではキーワード抽出のためのキーを名詞とする。キーワードとなる語は、キーワードリストに保管される（ステップS805）。

【0079】さらに、変数nがスコープを越えていないか判断する（ステップS806）。ここで言うスコープとは、検索文字列を示すノードからの距離がどれだけ離れたノードまでをキーワード抽出の対象とするかを表わす値とする。例えば、スコープを5とすると、検索文字列を示すノードからの距離が5以下のノードからキーワードを抽出することになる。本実施形態ではスコープを2とする。スコープを越えていない場合は、nを1増やし（ステップS807）、再度キーワードの検索を行う。越えている場合は、スコープ内の全ての語が検索されたので処理を終了し、処理結果であるキーワードリストは表示内容選択部603に渡される。キーワードが抽出されなかった場合は、空のキーワードリストが表示内容選択部603に渡される。

【0080】以上の処理により、例えば図14に示す構文解析結果が入力された場合には、「映画」がキーワードとして抽出され、表示内容選択部603に渡される。

【0081】以降は、第1の実施形態と同様に、表示内容選択部603によりキーワード抽出部602から渡されたキーワードリストとデータベース検索部104から渡された検索結果とに基づいて提示すべきデータが選択され、例えば図10に示すように出力部107により辞

書検索結果が表示される。

【0082】なお、本発明は、上記した第1、第2の実施形態に限定されるものではない。例えば、文解析としては形態素解析や構文解析を利用した例を上げたが、単純にキーワードを検索するだけでもかまわない。また、データベースとしては、上記では国語辞書を用いたが、その他の、例えば、外来語辞書、新語辞書、英和辞書などを用いてもよい。

【0083】（第3の実施形態）次に、本発明の第3の実施形態について説明する。

【0084】第1の実施形態では、指定された単語や熟語に基づいてデータベース106を検索して得た検索結果について、文解析結果から抽出したキーワードでさらに絞り込み等を行うものであったが、本実施形態では、文解析結果から抽出したキーワード（主たる検索条件）および指定された単語や熟語（従たる検索条件）に基づいてデータベース106に対する検索結果を得るようにしたものである。

【0085】本実施形態に係る情報検索装置の構成例は第1の実施形態と同様とする。すなわち、図1に示されるように、本情報検索装置は、入力部101、検索制御部102、文解析部103、データベース検索部104、解析知識105、データベース106、出力部107を備えている。

【0086】本実施形態では、検索制御部102における処理内容が第1の実施形態とは相違しているが、その他は基本的には第1の実施形態と同様である。

【0087】以下では、第1の実施形態と相違する点を中心に説明する。

【0088】入力部101から入力された文書データは、検索制御部102のメモリ（図示せず）に一旦記憶され、その後、入力部101を介して表示中の文書におけるある文字データ（単語や熟語）の検索が指示されると、文書データの所定の部分（例えば、指示された文字データを含む1文あるいは該1文を少なくとも含む複数文）が文解析部103へ渡される。

【0089】文解析部103は、検索制御部102の制御の下、該検索制御部102から文字データを受け取って解析知識105を利用して文解析を行い解析結果を出力する。文解析の方法としては、形態素解析、構文解析、意味解析など、一般的な手法を用いることができる。

【0090】検索制御部102は、入力部101から検索を指示された文字データと文解析部103から渡された文解析結果とを利用して、ユーザが求めているデータを得るための検索条件式を作成し、データベース検索部104に与え、検索を指示する。

【0091】データベース検索部104は、検索制御部102の制御の下、該検索制御部102から渡された検索条件式を受け取ってデータベース106を参照して検

素結果を出力する。

【0092】検索制御部102は、データベース検索部104によるデータベース106の検索結果を出力部107に渡す。

【0093】以下、具体例を用いながら本実施形態をより詳しく説明する。

【0094】ここでは文書は既に入力されているものとする。ユーザは入力部101を使い、文書内にある詳細情報を得たい文字列（すなわち検索文字列）を指示する。

【0095】具体例として、図16に例示するように、「…彼はジュラシックワールドの監督である。…。…」といった文書が表示された表示画面においてユーザがマウス等の入力装置を利用して「監督」という文字列部分をドラッグし、「監督」に関する詳細な情報が得たいと指示したとする。なお、「ジュラシックワールド」は映画名を表す固有名詞であるものとする。

【0096】検索制御部102は、位置情報等を利用しユーザの指示した検索文字列を含む1文を切り出し文解析部103に送る。

【0097】図16の例では、検索文字列「監督」を含む1文、「彼はジュラシックワールドの監督である。」が切り出され、文解析部103に送られる。

【0098】文解析部103は、検索制御部102より渡された1文に対して、解析知識105を利用して解析処理を行い、解析結果を検索制御部102に渡す。ここで言う解析処理とは、形態素解析、構文解析、意味解析等の全て、もしくはその一部指す。ここでは、第1の実施形態と同様、一例として、形態素解析を利用した場合について説明する。図17に、上記の1文すなわち「彼はジュラシックワールドの監督である。」に対して形態素解析を行った場合の解析結果の一例を示す。

【0099】検索制御部102は、文解析部103からの解析結果および検索文字列を利用して検索条件式を作成しデータベース検索部104に渡す。

【0100】ここで、図18に本実施形態における検索制御部102の構成例を示す。図18に示されるように、この検索制御部102は、制御記憶部2501、検索条件式作成部2502から構成される。

【0101】制御記憶部2501は、入力部101から入力された文書を記憶したり、ユーザからの検索指示を受け取り、1文を切り出し文解析部103に送ったり、データベースの検索結果を出力部107に送ったりといった、検索全体に関わる制御を行う。

【0102】検索条件式作成部2502は、文解析部103からの解析結果の中から、辞書を検索するために有効なキーワードを抽出し、検索条件式を作成してデータベース検索部104に渡す。

【0103】図19に、検索条件式作成部2502による検索条件式の作成手順の一例を示す。

【0104】まず、変数nに1をセットし（ステップS2602）、検索文字列（図17の場合、「監督」）からのn語離れた語（距離）を抽出する（ステップS2603）。

【0105】形態素解析結果は、一語一語キーワードになるかどうか判断される（ステップS2604）。ここで言うキーワードは、検索条件式を作成するために有効なもので、名詞であるか、固有名詞であるかといった品詞情報や、コンピュータ用語、経済用語などといった分野情報等をキーとして抽出された用語を指す。ここではキーワード抽出のためのキーを固有名詞とする。キーワードとなる語はキーワードリストに保管される（ステップS2605）。

【0106】さらに、変数nがスコープを越えていないか判断する（ステップS2606）。ここで言うスコープとは、検索文字列から何語離れた語までをキーワード抽出の対象とするかを表わす値とする。例えば、スコープを5とすると、検索文字列の前後5単語からキーワードを抽出することになる。本実施形態ではスコープを3とする。スコープを越えていない場合は、nを1増やし（ステップS2607）、再度キーワードの検索を行う。越えている場合は、キーワードリストと検索文字列から検索条件式を作成し（ステップS2608）、処理を終了し（ステップS2609）、検索条件式がデータベース検索部104に渡される。キーワードリストと検索文字列から検索条件式を作成する手法については、公知の手法を用いれば良い。

【0107】以上の処理により、例えば図17で示す形態素解析結果が入力された場合には、「ジュラシックワールド」がキーワードとして抽出され、検索文字列「監督」と共に「ジュラシックワールドの監督」を検索するための検索条件式が図20に示すように作成される。

【0108】データベース検索部104は、検索制御部102により渡された検索条件式を使ってデータベースの中から対象となるデータを取り出し検索制御部102に渡す。

【0109】例えば、データベース106が図21に例示する映画事典のようなものであるとすると、検索制御部102から渡された図20に示す検索条件式より「ジュラシックワールド」に関する情報を取り出し、条件式からその素性として監督が要求されているので、「スティーブ・○○○○○」という文字列が取り出される。

【0110】ここで、図21のデータの“< >”で囲まれた部分はタグである。例えば、“<見出し語>”から、“<\見出し語>”で挟まれた部分が事典の見出し語となる。この例では事典の見出し語は“ジュラシックワールド”となる。次に見出し語に対する公開年度が記述されている。ここでは、1993年に公開されたことを示す。次に見出し語の製作者、監督と記述されている。以下記述内容の説明は省略する。なお、「スティー

ブ・○○○○○」等は人名を表しているものとする。

【0111】検索された文字列は、検索制御部102を介して出力部107に渡され、辞書検索結果が表示される。

【0112】例えば、図22に示すように、出力部107の出力装置により辞書検索結果が、ユーザに呈示される。

【0113】なお、上記では、キーワードリストと検索文字列とから検索条件式を作成し、この検索条件式によりデータベースを検索して得た検索結果を提示するようにしたが、その代わりに、キーワードリストから検索条件式を作成し、この検索条件式によりデータベースを検索した後に、検索結果のうちから検索文字列に該当するものを選択して提示するようにしてもよい。

【0114】(第4の実施形態)次に、本発明の第4の実施形態について説明する。

【0115】第3の実施形態では、文解析部103の解析結果として形態素解析を利用する場合について説明したが、本実施形態では、文解析部103の解析結果として構文解析を利用する場合について説明する。

【0116】なお、以下では、第3の実施形態と相違する点を中心に説明する。

【0117】まず、ユーザにより検索指示が入力されると、第1の実施形態と同様に、例えば検索文字列「監督」を含む1文すなわち「彼はジュラシックワールドの監督である。」が切り出され、文解析部103に送られる。

【0118】ここで、図23に、文解析部103の解析結果として形態素解析・構文解析を利用した場合の出力結果の一例を示す。文解析部103の解析結果が、図23で表わすような構文解析結果(概念依存構造データ)の場合、検索制御部102の検索条件式作成部2502を変更することにより構文解析結果を利用した場合の検索装置を実現できる。

【0119】図24に構文解析結果を利用する場合における検索制御部102内の検索条件式作成部2502の処理手順の一例を示す。

【0120】まず、変数nに1をセットする(ステップS2702)。

【0121】検索文字列から何回リンクを辿ったかという値nを抽出する(ステップS2703)。この値nは検索文字列を示すノードからの距離を表す。すなわち、図23の「監督」のノードから見ると「ジュラシックワールド」「である」のノードは距離が1で「彼」のノードは距離が2となる。

【0122】構文解析結果は、一語一語キーワードになるかどうか判断される(ステップS2704)。ここで言うキーワードは、検索結果を絞り込むためのもので、名詞であるか、固有名詞であるかといった品詞情報や、コンピュータ用語、経済用語などといった分野情報等を

キーとして抽出された用語を指す。ここではキーワード抽出のためのキーを固有名詞とする。キーワードとなる語は、キーワードリストに保管される(ステップS2705)。

【0123】さらに、変数nがスコープを越えていないか判断する(ステップS2706)。ここで言うスコープとは、検索文字列を示すノードからの距離がどれだけ離れたノードまでをキーワード抽出の対象とするかを表わす値とする。例えば、スコープを5とすると、検索文字列を示すノードからの距離が5以下のノードからキーワードを抽出することになる。本実施形態ではスコープを2とする。スコープを越えていない場合は、nを1増やし(ステップS2707)、再度キーワードの検索を行う。越えている場合は、スコープ内の全ての語が検索されたので、キーワードリストと検索文字列から検索条件式を作成し(ステップS2708)、処理を終了(ステップS2709)、検索条件式がデータベース検索部104に渡される。キーワードリストと検索文字列から検索条件式を作成する手法については現在の技術で一般的な手法を用いればよい。

【0124】以上の処理により、例えば図23で示す構文解析結果が入力された場合には、「ジュラシックワールド」がキーワードとして抽出され、検索文字列「監督」と共に「ジュラシックワールドの監督」を検索するための検索条件式が図20に示すように作成される。

【0125】以降は、第3の実施形態と同様に、データベース検索部104により上記の検索条件式を使ってデータベースの中から対象となるデータが取り出され、例えば図22に示すように出力部107の出力装置により辞書検索結果がユーザに呈示される。

【0126】なお、本発明は、上記した第3、第4の実施形態に限定されるものではない。例えば、文解析は形態素解析や構文解析を利用した例を上げたが、単純にキーワードを検索するだけでもかまわない。また、データベースとしては、上記では国語辞書を用いたが、その他の、例えば、国語辞書、外来語辞書、新語辞書、英和辞書などを用いても構わない。

【0127】また、第1～第4の実施形態では、キーワード抽出のための文解析の対象として、ユーザの指示した検索文字列を含む1文を切出したが、この文解析の対象とする文の切り出し方については、この他にも種々の方法が考えられ、例えば、ユーザの指示した検索文字列を含む1文とその前後複数の文を一纏まりとして切り出すようにしてもよい。

【0128】(第5の実施形態)次に、本発明の第5の実施形態について説明する。

【0129】第1～第4の実施形態では、文解析結果から得たキーワードをそのまま検索もしくは検索結果の選択に用いたが、本実施形態は、文解析結果から得たキーワードに基づいてさらに別の1または複数のキーワード

を求め、前者のキーワードの代わりにあるいは前者のキーワードとともに、後者のキーワードを用いるようにするものである。

【0130】この構成は第1～第4の実施形態のいずれにも適用可能であるが、以下では、前述した第1の実施形態において、文解析結果から抽出されたキーワードから所定のデータベースを用いて他の関連するキーワードを抽出し、それらのキーワードのうち少なくとも一つを用いて、検索結果から特定のデータを選択する場合について説明する。

【0131】以下では、第1の実施形態と相違する点を中心に説明する。

【0132】本実施形態に係る情報検索装置の構成例は第1の実施形態と同様とする。すなわち、図1に示されるように、本情報検索装置は、入力部101、検索制御部102、文解析部103、データベース検索部104、解析知識105、データベース106、出力部107を備えている。

【0133】本実施形態では、検索制御部102に上記のような機能が付加される点が第1の実施形態とは相違しているが、その他は基本的には第1の実施形態と同様である。

【0134】以下、具体例を用いながら本実施形態をより詳しく説明する。

【0135】ここでは、文書は既に入力されているものとする。ユーザは入力部101を使い、文書内にある詳細情報が得たい文字列（すなわち検索文字列）を指示する。

【0136】具体例として、図16に例示するように、「…彼はジュラシックワールドの監督である。…」と…。」といった文書が表示された表示画面においてユーザがマウス等の入力装置を利用して「監督」という文字列部分をドラッグし、「監督」に関する詳細な情報が得たいと指示したとする。なお、「ジュラシックワールド」は映画名を表す固有名詞であるものとする。

【0137】検索制御部102は、位置情報等を利用してユーザの指示した検索文字列を含む一文を切り出し、文解析部103に送る。

【0138】図16の例では、検索文字列「監督」を含む一文、「彼はジュラシックワールドの監督である。」が切り出され、文解析部103に送られる。

【0139】文解析部103は、検索制御部102より渡された1文に対して解析知識105を利用して解析処理を行い、解析結果を検索制御部102に渡す。ここで言う解析処理とは、形態素解析、構文解析、意味解析等の全て、もしくはその一部指す。ここでは、第1の実施形態と同様、一例として、形態素解析を利用した場合について説明する。図17に、上記の1文すなわち「彼はジュラシックワールドの監督である。」に対して形態素解析を行った場合の解析結果の一例を示す。

【0140】データベース検索部104は、検索制御部102から入力された検索文字列「監督」を使ってデータベース106の中から対象となるデータを取り出し検索制御部102に渡す。

【0141】ここでは、第1の実施形態と同様に、データベース106が国語辞書のようなものであり、図4に例示するような結果が取り出されたものとする。

【0142】文解析部103からの解析結果およびデータベース検索部104からの検索結果は、検索制御部102に返され、ユーザに表示する内容が選択される。

【0143】ここで、図25に本実施形態における検索制御部102の構成例を示す。図25に示されるように、この検索制御部102は、制御記憶部2401、キーワード抽出部2402、データベース2403、表示内容選択部2404から構成される。

【0144】制御記憶部2401は、第1の実施形態と同様であり、入力部101から入力された文書を記憶したり、ユーザからの検索指示を受け取りデータベース検索部104に送ったり、1文を切り出し文解析部103に送るといった、検索全体に関わる制御を行う。

【0145】キーワード抽出部2402は、文解析部103から渡された解析結果の中からデータベース2403を参照しつつ、提示すべき内容を選択するのに有効なキーワードを抽出し、キーワードリストを作成して制御記憶部2401に渡す。

【0146】本実施形態で使用されるデータベース2403は、語の概念の上位下位関係を表わすシソーラスデータとする。図26にシソーラスデータの一例を示す。図26では、「ジュラシックワールド」「デイハード」の上位概念として「アクション」があり、その上位概念として「映画」があるといった上位下位関係が示されている。なお、「デイハード」は映画名を表す固有名詞、「アクション」は映画のジャンルを表す一般名詞であるものとする。

【0147】図27にキーワード抽出部2402のキーワードリスト作成の処理手順の一例を示す。

【0148】まず、変数Nに1をセットする（ステップS2902）。この値は検索文字列（図17の場合、「監督」）からの距離を表わす（ステップS2903）。ここで言う距離とは、検索文字列から何語離れているかという値である。

【0149】形態素解析結果は、一語一語キーワードになるかどうか判断される（ステップS2903、S2904）。ここで言うキーワードは、検索結果を絞り込むためのもので、名詞であるか、固有名詞であるかといった品詞情報や、コンピュータ用語、経済用語といった分野情報等をキーとして抽出された用語を指す。ここではキーワード抽出のためのキーを固有名詞とする。キーワードとなる語はキーワードリストに保管される（ステップS2911）。

【0150】なお、本具体例では、図17に示す形態素解析結果が入力された場合、変数N=2のときに、「ジュラシックワールド」がキーワードとして抽出される。

【0151】キーワードが抽出されたならば、次に、キーワードに関連する用語を抽出する処理を行う（ステップS2912～S2917）。

【0152】ここではデータベースとしてシソーラスを用いそのシソーラス上のキーワードとのノードとの距離によってキーワード候補を抽出する。ここで言うシソーラス上の距離とは、キーワードのノードから何回リンクを辿ったかという値である。すなわち、図26において、「ジュラシックワールド」のノードから見ると、「アクション」のノードは距離が1で「デイハード」「映画」のノードは距離が2となる。

【0153】まず、変数Mに1をセットする（ステップS2912）。この値はステップS2904で抽出されたキーワードを示すノードからのシソーラス上の距離を表わす。シソーラス上の語は一語一語キーワードとなるか判断される（ステップS2913、S2914）。シソーラスを使ってキーワード候補を捜すためのキーは、形態素解析結果を使ってキーワード候補を捜すためのキーと異なっていておかまわない。ここではキーを一般名詞とする。キーワードとなる語はキーワードリストに保管される（ステップS2915）。

【0154】さらに、変数Mがスコープを越えていないか判断する（ステップS2916）。ここで言うスコープとは、キーワードを表わすノードからのシソーラス上の距離がどれだけ離れたノードまでをキーワード抽出の対象とするかを表す値とする。例えば、スコープを5とすると、キーワードを表わすノードからのシソーラス上の距離が5以下のノードからキーワードを抽出することになる。本実施形態ではスコープを2とする。スコープを越えていない場合は、Mを1増し（ステップS2917）、再度キーワードの検索を行う。越えている場合は、シソーラス中のスコープ内の全ての語が検索されたので、ステップS2906に戻る。

【0155】本具体例では、キーワードとして「ジュラシックワールド」が抽出された場合、該キーワードから距離が2以下でありかつ一般名詞であるノードとして「アクション」「映画」が抽出される。

【0156】ステップS2906では、変数Nがスコープを越えていないか判断する。ここで言うスコープとは、検索文字列から何語離れた語までをキーワード抽出の対象とするかを表わす値とする。例えば、スコープを5とすると、検索文字列の前後5単語からキーワードを抽出することになる。本実施形態ではスコープを3とする。スコープを越えていない場合は、ステップS2907でNを1増やし、再度キーワードの検索を行う。越えている場合は、スコープ内の全ての語が検索されたので処理を終了し、処理結果であるキーワードリストは表示

内容選択部に渡される。キーワードが抽出されなかった場合は、空のキーワードリストが表示内容選択部に渡される。

【0157】以上の処理により、例えば図17に示す形態素解析結果が入力された場合には、変数N=2のときに、「ジュラシックワールド」「アクション」「映画」がキーワードとして抽出され、表示内容選択部2404に渡される。

【0158】表示内容選択部2404は、第1の実施形態で示した手法を用いて、検索結果の語義に対して評価点を計算し、その点数の高いものを出力部107に渡す。本実施形態のキーワード数テーブルを図28に示す。この結果、最も適切な語義として二番目の語義データが選ばれる。

【0159】そして、例えば、図29に示すように、出力部107により辞書検索結果が表示される。

【0160】なお、上記では図25のデータベース2403としてシソーラスを用いたが、これに限定されず、種々のデータベースを用いることが可能である。

【0161】また、例えば、第1、第2の実施形態において、入力文書を和文とし、図1の検索対象とするデータベース106として和英辞書を用いる場合に、検索結果として英文データが得られ、日本語のキーワードでは検索結果の選択ができないようなケースであっても、データベース2403として和英辞書を用いてキーワードを英語の訳語に変換することで、検索結果から提示すべきデータを選択することができるようになる。もちろん、他の言語の場合も同様である。

【0162】また、例えば、データベース2403として和英（英和）辞書を用いてキーワードを訳語に変換することで、訳語を使用してデータベース検索部104でデータベースを検索することができる。

【0163】また、上記では、文解析により得られたキーワードの上位概念をシソーラスを用いて求めたが（これを1回のキーワードの変換とする）、例えば、まずシソーラスを用いてキーワードの上位概念を求め、次に和英辞書を用いてその上位概念の訳語を求め、これを検索等に用いる、というように、2回以上のキーワードの変換を行うようにしてもよい。

【0164】また、例えば、データベース2403として相異なる複数の辞書を用い、各辞書によりそれぞれ得られたキーワードを併せて用いるようにしてもよい。

【0165】また、最初に第1～第4の実施形態のように文解析結果から得たキーワードをそのまま用いて検索処理を行い、当該キーワードが有効でなかった場合に（例えば、当該キーワードに一致する語句が検索結果や見出し語に存在しなかった場合）、本実施形態のように文解析結果から得たキーワードに基づいて得た別の1または複数のキーワードを用いて検索処理を行うようにしてもよい。

【0166】(第6の実施形態)第1～第5の実施形態では、検索結果として提示する情報の形態として、文字データである場合について説明したが、本実施形態では、第1～第5の実施形態において、該形態として、写真、図、音声、画像、テキスト等といった複数のマルチメディア情報を取り扱う場合について説明する。

【0167】検索結果として複数のマルチメディア情報を取り扱うようにする場合において、任意の検索文字列に対してすべてのマルチメディア情報が表示(もしくは再生等)可能というわけではない。例えば、音楽のタイトルが文中に出てきた場合に、音声データは出力可能であるが、音楽であるため写真等は出力できない、といった状況が考えられる。

【0168】そこで、任意の検索文字列が指定された場合、その検索文字列で表示可能な情報の種類をポップアップメニュー等を使用して、ユーザに提示し、ユーザに表示する情報を選択させるようにすると好ましい。以下では、そのいくつかの例を示す。

【0169】ここで、図16の例のような文書すなわち「…彼はジュラシックワールドの監督である。…」がすでに入力されているものとする。

【0170】ユーザは入力部101を使い文書内にある詳細情報を得たい文字列(すなわち検索文字列)を指示する。

【0171】例えば、「監督」という文字列の詳細情報が得たいとすると、図16に例示するように、ユーザはマウス等の入力装置を利用して「監督」という文字列部分をドラッグする。すると、検索制御部は、これにตอบสนองして、例えば図30に示すように、表示可能なマルチメディア情報をポップアップメニューでユーザに提示する。

【0172】次に、ユーザは、そのポップアップメニューから自分が得たい詳細情報の種類を選択する。例えば、「監督」の「作品」に関する詳細情報が得たい場合、図30に示すようにマウス等の入力装置を利用して「作品」の部分を選ぶことにより指示する。

【0173】当該情報検索装置では、すでに説明したような手法を使って、例えば図31に示すように、ユーザの指示した「監督」の「作品」のリストを表示する。なお、「映画A」、「映画B」、「映画C」および「映画D」は、映画名を表す固有名詞であるものとする。

【0174】また、図30のポップアップメニューで「語義」が選択指示されると、例えば図32に示すように、「監督」の語義が表示される。

【0175】なお、上記の指定は、ファンクションキー、マウスのクリック回数、指定アイコンのドラッグアンドドロップ等で行ってもよい。

【0176】また、上記のように情報の種類を選択してからデータを検索するのではなく、予めデータを検索してから表示等できる情報を提示し、ユーザに指示された

種類に該当する検索結果を表示等するようにしてもよい。

【0177】以下では、以上の各実施形態における種々のバリエーションについて説明する。

【0178】上記では、検索制御部の制御の下で該検索制御部を介して各部にデータを受け渡しするものとして説明したが、データ受け渡しにあっては、実データを渡すようにしてもよいし、その代わりに、実データが格納されているアドレスなどのポインタ情報を渡すようにしてもよい。

【0179】また、上記では、データベース106は当該情報検索装置に備え付けられたものとして説明したが、データベース106は当該情報検索装置では備えず、当該情報検索装置が外部のデータベース106にアクセスするようにしてもよい。また、いずれの場合においても、データベース106として複数種類のものを検索可能とし、ユーザが検索対象とする1または複数のデータベースを指定可能としてもよい。

【0180】また、同様に、第5の実施形態では、データベース2403は検索制御部102内に備え付けられたものとして説明したが、データベース2403は当該情報検索装置では備えず、当該情報検索装置が外部のデータベース2403にアクセスするようにしてもよい。

【0181】また、検索結果を表示する際に、検索等に使用したキーワードを併せて提示するようにしてもよい。

【0182】また、上記した各実施形態の検索機能は、適宜組み合わせることも可能である。

【0183】また、以上の各実施形態において説明した機能は、ソフトウェアによっても実現可能であり、当該情報検索装置は、専用機として構成することも可能であるが、汎用の計算機を利用して構成することも可能である。

【0184】また、以上の各実施形態において説明した機能は、記録媒体に記憶させたものとして、すなわちコンピュータに所定の手順を実行させるための(あるいはコンピュータに所定の手段として機能させるための)プログラムを記録したコンピュータ読取り可能な記録媒体として実施することもできる。例えば、図1において、検索制御部102、文解析部103、データベース検索部104を、検索制御機能、文解析機能、データベース検索機能として情報検索のための記録媒体に展開することで実施できる。また、この場合、解析知識105やデータベース106は、当該記録媒体に記憶して実現するようにしてもよいし、あるいは当該記録媒体とは別の記憶媒体(メモリ)に記憶して、検索制御機能、文解析機能、データベース検索機能が記憶された当該記録媒体と連携して可動するようにしてもよい。

【0185】また、各実施形態における検索機能は、例

えば、文書を作成もしくは編集するためのソフト、WWWや電子メールにより取得した文書を閲覧するためのソフト、機械翻訳ソフトなどに組み込む形で実施することも可能である。また、例えばネットワークを介するなどして外部の要求元から文書データと指定の検索文字列を受信し、検索結果を該要求元へ送信するような検索サービスシステムとして実施することも可能である。

【0186】本発明は、上述した実施の形態に限定されるものではなく、その技術的範囲において種々変形して実施することができる。

【0187】

【発明の効果】本発明によれば、ユーザは必要な情報を効率良く得ることができ、文章の理解を早めることに効果がある。つまり、ユーザは必要な情報を自分で絞り込むことなく情報が呈示されるため、ユーザの負荷が軽減されるのである。

【図面の簡単な説明】

【図1】本発明の実施の形態に係る情報検索装置の構成例を示す図

【図2】文書の表示例を示す図

【図3】対象となった文の形態素解析結果の一例を示す

【図4】国語辞書検索結果の一例を示す図

【図5】検索制御部の構成例を示す図

【図6】形態素解析結果を利用したキーワード抽出処理の手順の一例を示すフローチャート

【図7】表示情報を選択するための処理手順の一例を示すフローチャート

【図8】キーワード数テーブル（初期状態）の一例を示す図

【図9】キーワード数テーブル（終了状態）の一例を示す図

【図10】検索結果の表示例を示す図

【図11】検索結果の表示例を示す図

【図12】検索結果の表示例を示す図

【図13】検索結果の表示例を示す図

【図14】対象となった文の構文解析結果の一例を示す図

【図15】構文解析結果を利用したキーワード抽出処理の手順の一例を示すフローチャート

【図16】文書の表示例を示す図

*40

*【図17】対象となった文の形態素解析結果の一例を示す図

【図18】検索制御部の構成例を示す図

【図19】形態素解析結果を利用したキーワード抽出処理の手順の一例を示すフローチャート

【図20】検索条件式の一例を示す図

【図21】映画事典検索結果の一例を示す図

【図22】検索結果の表示例を示す図

【図23】対象となった文の構文解析結果の一例を示す

10 図

【図24】構文解析結果を利用したキーワード抽出処理の手順の一例を示すフローチャート

【図25】検索制御部の構成例を示す図

【図26】シソーラスデータの一例を示す図

【図27】表示情報を選択するための処理手順の一例を示すフローチャート

【図28】キーワード数テーブル（終了状態）の一例を示す図

【図29】検索結果の表示例を示す図

20 【図30】ポップアップメニューの表示例を示す図

【図31】検索結果の表示例を示す図

【図32】検索結果の表示例を示す図

【符号の説明】

101…入力部

102…検索制御部

103…文解析部

104…データベース検索部

105…解析知識

106…データベース

107…出力部

601…制御記憶部

602…キーワード抽出部

603…表示内容選択部

2501…制御記憶部

2502…検索条件式作成部

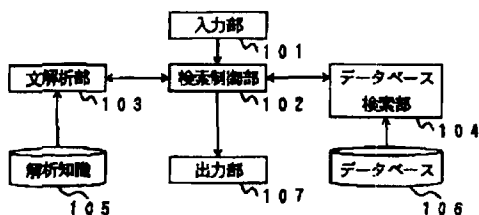
2401…制御記憶部

2402…キーワード抽出部

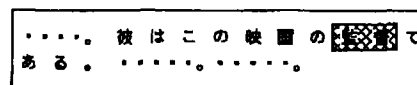
2403…データベース

2404…表示内容選択部

【図1】



【図2】



【図3】

辞 順	見出し語	品 詞
0	彼	人称代名詞
1	は	助詞
2	この	代名詞
3	映画	名詞
4	の	助詞
5	監督	名詞
6	である	助詞

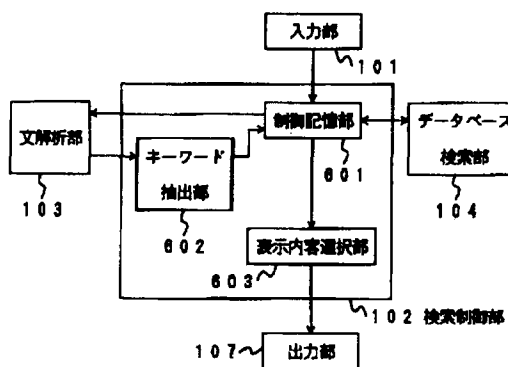
【図4】

<見出し語>監督<見出し語>
 <品詞情報>名・他動詞<品詞情報>
 <語義部1>人の上に立って指揮すること、人、機関。
 「部下をーする・工事の現場ー」<語義部1>
 <語義部2>映画・演劇などで演出をする人。
 「演劇の舞台ー・映画ー」<語義部2>
 <語義部3>……………<語義部3>
 ……
 ……

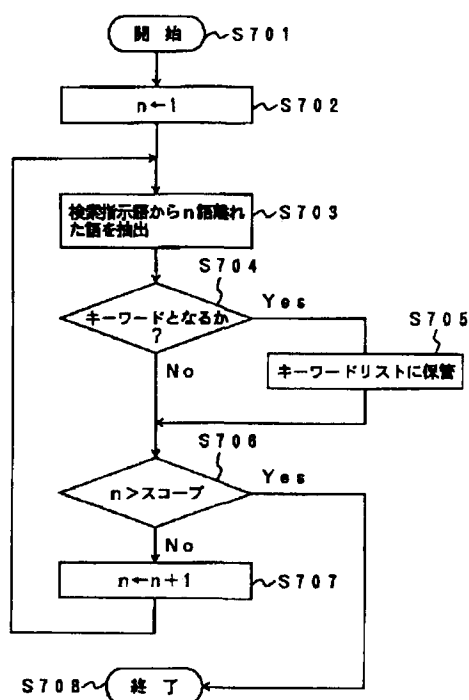
【図13】

制作者
 ……。彼は、この映画の[監督]
 である。……………

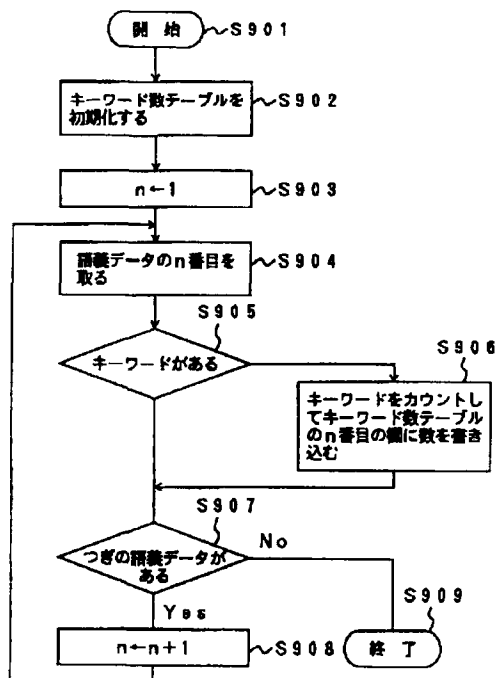
【図5】



【図6】



【図7】



【図8】

エントリー番号	キーワード出現回数	語義データ
1	0	語義部1
2	0	語義部2
3	0	語義部3

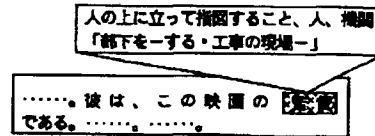
【図9】

エントリー番号	キーワード出現回数	語義データ
1	0	語義部1
2	2	語義部2
3	0	語義部3

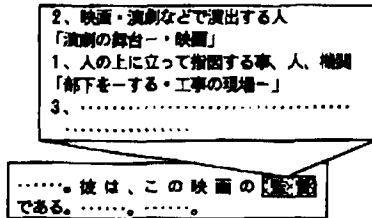
【図10】



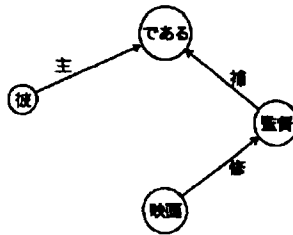
【図11】



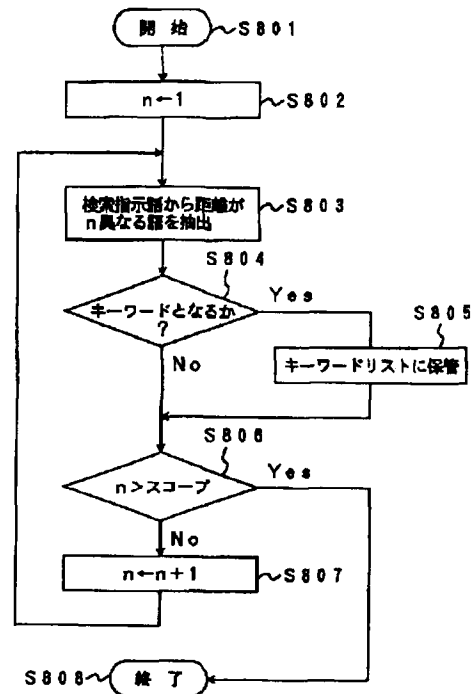
【図12】



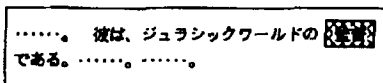
【図14】



【図15】



【図16】



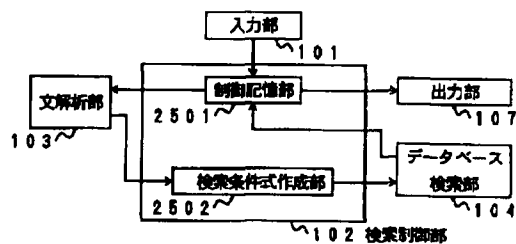
【図20】

search ((見出し語: ジュラシックワールド) (属性: 監督))

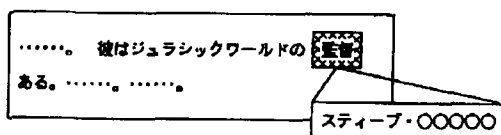
【図17】

順 取	見出し語	品 詞
0	彼	人称代名詞
1	は	助詞
2	ジュラシックワールド	固有名詞
3	の	助詞
4	監督	名詞
5	である	動詞

【図18】



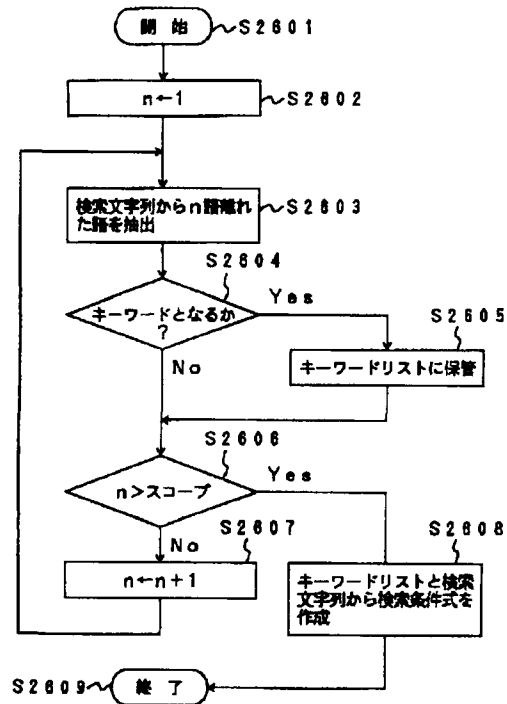
【図22】



【図28】

エントリー番号	キーワード出現回数	照会データ
1	0	照会部1
2	2	照会部2
3	0	照会部3

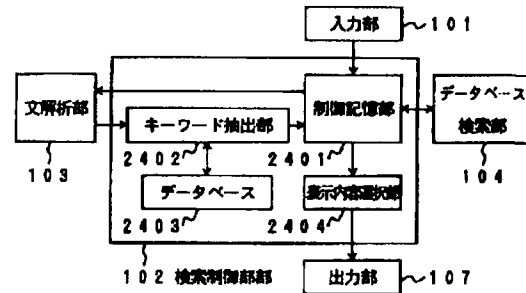
【図19】



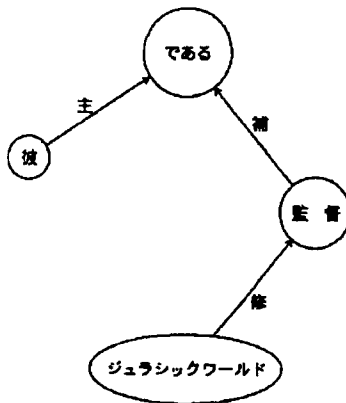
【図21】

<見出し語>ジュラシックワールド<\見出し語>
 <公開年度>1993<\公開年度>
 <制作>ジャクリーン・〇〇〇〇〇〇<\制作>
 <監督>スティーブ・〇〇〇〇〇〇<\監督>
 <出演1>トム・〇〇〇〇〇〇<\出演1>
 <出演2>チャーリー・〇〇〇〇〇〇<\出演2>
 <解説>近未来。遺伝子工学を駆使して再現された恐竜の.....<\解説>

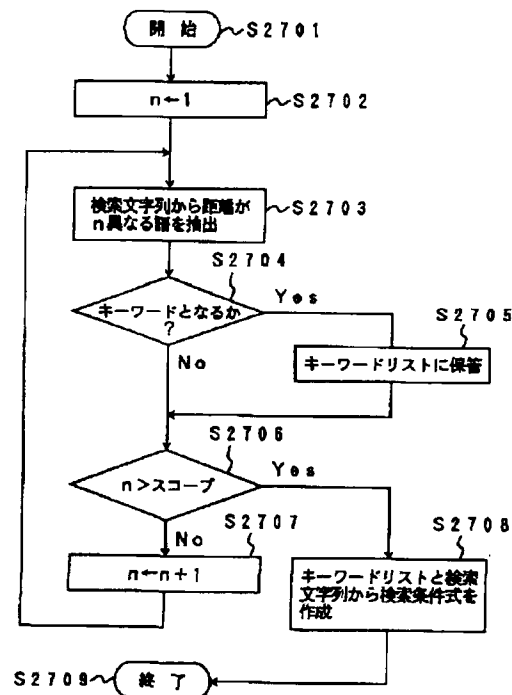
【図25】



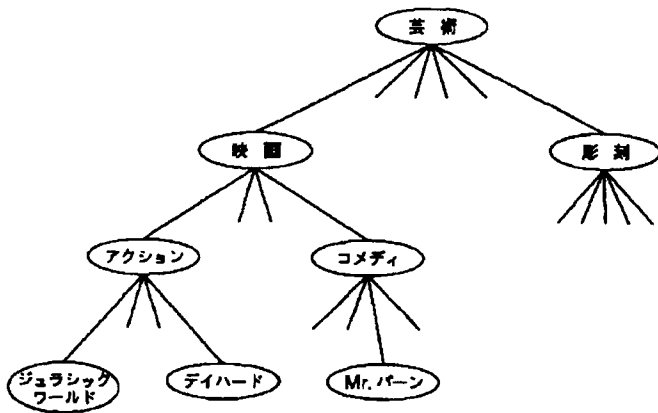
【図23】



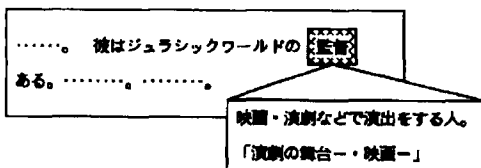
【図24】



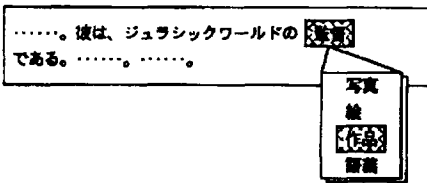
【図26】



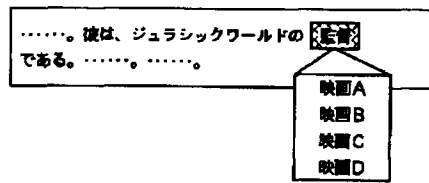
【図29】



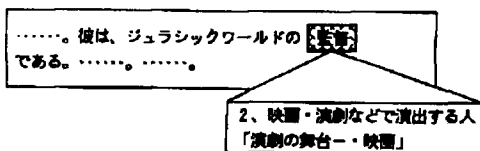
【図30】



【図31】



【図32】



【図27】

